PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-356939
(43)Date of publication of application : 26.12.2001

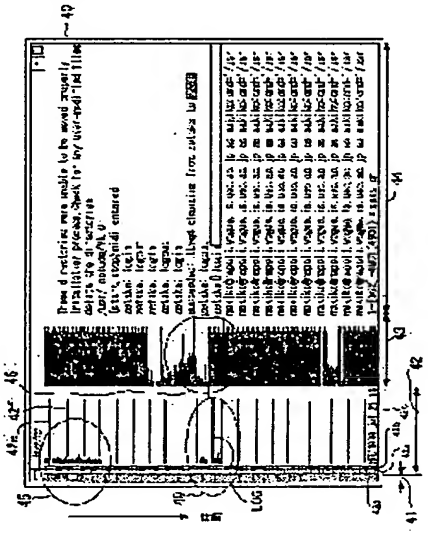
(51)Int.Cl. G06F 11/34

(21)Application number : 2000-177205
(22)Date of filing : 13.06.2000

(71)Applicant : TOKYO ELECTRIC POWER CO INC:THE
(72)Inventor : TAKADA TETSUJI
KOIKE HIDEKI
AMARI HARUO
OKADA MIKIO

(54) LOG INFORMATION ANALYZER, ITS METHOD AND RECORDING MEDIUM

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a log information analyzer, its method and a recording medium capable of reducing time to be required for an analyzing work of log information, integrating pieces of log information with various data forms and properties of uneven distribution on a file system, extracting log information to indicate an abnormality which is not the known one and enabling a system manager to exactly grasp noticeable log information among a vast quantity of log information by characters.
SOLUTION: Pieces of log information unevenly distributed in the log file are integrated on the basis of time information 17 by converting pieces of the log information recorded by various data forms into intermediate forms as multipurpose log formats. The log information to indicate the abnormality different from the known one is extracted by calculating appearance frequency 34 by word and appearance frequency 37 by phrase formed by linking two words. The time to be required for the analyzing work of the log information is reduced since the log information is combined with extracted feature information, defined as a graphic and offered to the system manager.



LEGAL STATUS

[Date of request for examination] 09.11.2000
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]
[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（Ａ）　(11)特許出願公開番号
特開2001－356939
（P2001－356939A）
(43)公開日　平成13年12月26日(2001.12.26)

(51)Int.Cl.⁷	識別記号	ＦＩ	ﾃｰﾏｺｰﾄﾞ(参考)
Ｇ０６Ｆ　11/34		Ｇ０６Ｆ　11/34	Ｃ　５Ｂ０４２

審査請求　有　　請求項の数17　ＯＬ　（全 15 頁）

(21)出願番号　特願2000－177205(P2000－177205)

(22)出願日　平成12年６月13日(2000.6.13)

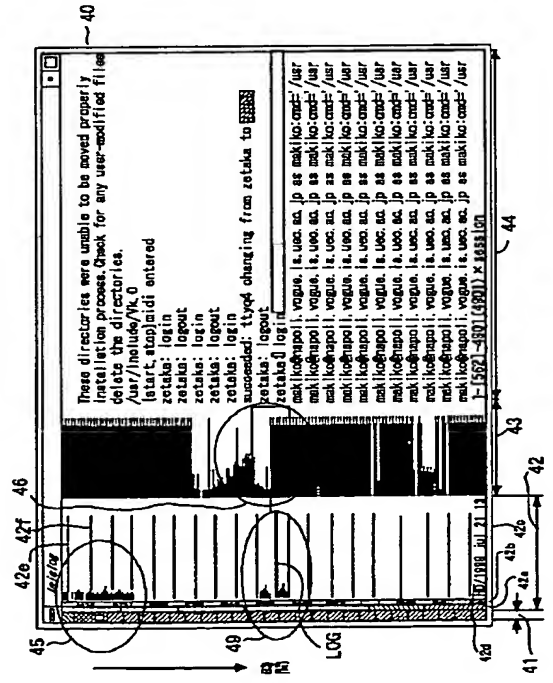
特許法第30条第１項適用申請有り　1999年12月15日　インターネットコンファレンス'99実行委員会発行の「インターネットコンファレンス'99論文集」に発表

(71)出願人　000003687
東京電力株式会社
東京都千代田区内幸町１丁目１番３号
(72)発明者　高田　哲司
東京都調布市飛田給２丁目46番12号　エミールマンション205
(72)発明者　小池　英樹
東京都杉並区荻窪５丁目29番17号　1104
(72)発明者　甘利　治雄
神奈川県横浜市鶴見区江ヶ崎町４番１号
東京電力株式会社システム研究所内
(74)代理人　100082175
弁理士　高田　守　（外２名）

最終頁に続く

(54)【発明の名称】　ログ情報解析装置、方法および記録媒体

(57)【要約】
【課題】　ログ情報の解析作業に要する時間を減少させ、種々のデータ形式とファイル・システム上の偏在性とを有するログ情報を統合し、既知の異常ではない異常を示すログ情報を抽出することができ、システム管理者が文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することができるログ情報解析装置、方法および記録媒体を提供する。
【解決手段】　種々のデータ形式により記録されているログ情報を汎用ログフォーマットという中間形式に変換することにより、ログファイル中に偏在していたログ情報を時刻情報１７を基準として統合することができる。単語別出現頻度３４と二単語連結句別出現頻度３７とを求めることにより、既知の異常とは異なる異常を示すログ情報を抽出することができる。ログ情報と抽出された特徴情報とを組み合わせて図形化してシステム管理者に提示することができるため、ログ情報の解析作業に要する時間を減少させることができる。



【特許請求の範囲】
【請求項１】　コンピュータの稼動状況を示すログ情報が記録されたログ情報記録部と、
前記ログ情報記録部に記録されたログ情報から該ログ情報が記録された時刻と所定の文字列とを抽出し、所定の時刻形式に変換された時刻と所定のフォーマットに変換された所定の文字列とを含む汎用ログを作成する汎用ログ作成手段と、
前記汎用ログ作成手段により作成された汎用ログを、前記所定の時刻形式に変換された時刻の順に記録する汎用ログ記録手段と、
前記汎用ログ記録手段により記録された汎用ログを有する汎用ログ記録部と、
前記汎用ログ記録部に記録された汎用ログに対応するログ情報の長さと該ログ情報中に現れる語の出現頻度とを含む特徴情報を前記ログ情報記録部から抽出する特徴情報抽出手段と、
前記特徴情報抽出手段により抽出されたログ情報を図形化して前記ログ情報と共に前記所定の時刻形式に変換された時刻の順に表示する表示手段とを備えたことを特徴とするログ情報解析装置。
【請求項２】　請求項１記載のログ情報解析装置において、前記特徴情報に含まれる語の出現頻度は、前記所定のフォーマットに変換された所定の文字列中に１単語が単独で出現する頻度と少なくとも２単語が連続して出現する頻度とを用いて求められることを特徴とするログ情報解析装置。
【請求項３】　請求項１記載のログ情報解析装置において、前記表示手段が表示するログ情報は、前記特徴情報に含まれる語の出現頻度が所定の値を有する語または指定された所定の語を少なくとも含む文字列が識別可能に表示されることを特徴とするログ情報解析装置。
【請求項４】　請求項１ないし３のいずれかに記載のログ情報解析装置において、前記表示手段は、前記所定のフォーマットに変換された所定の文字列が所定の種類に該当するログ情報を選択して表示することを特徴とするログ情報解析装置。
【請求項５】　請求項１ないし３のいずれかに記載のログ情報解析装置において、前記表示手段は、前記特徴情報に含まれる語の出現頻度が所定の範囲内にある語を含むログ情報を選択して表示することを特徴とするログ情報解析装置。
【請求項６】　請求項１ないし３のいずれかに記載のログ情報解析装置において、前記表示手段は、前記ログ情報の長さが所定の範囲内にあるログ情報を選択して表示することを特徴とするログ情報解析装置。
【請求項７】　請求項１ないし３のいずれかに記載のログ情報解析装置において、前記表示手段は、前記所定の時刻形式に変換された時刻が所定の範囲内にあるログ情報を選択して表示することを特徴とするログ情報解析装置。
【請求項８】　請求項１ないし３のいずれかに記載のログ情報解析装置において、前記表示手段は、所定の文字列を有するログ情報を選択して表示することを特徴とするログ情報解析装置。
【請求項９】　請求項１ないし８のいずれかに記載のログ情報解析装置において、前記汎用ログ作成手段における前記所定の時刻形式は、所定の基準となる時刻から前記ログ情報が記録された時刻までの経過時間であることを特徴とするログ情報解析装置。
【請求項１０】　請求項１ないし９のいずれかに記載のログ情報解析装置において、前記表示手段における前記ログ情報の図形化は、所定の時間あたりの該ログ情報の数のヒストグラム化であることを特徴とするログ情報解析装置。
【請求項１１】　請求項１ないし９のいずれかに記載のログ情報解析装置において、前記表示手段における前記ログ情報の図形化は、該ログ情報に含まれる所定の文字列の長さに対応した線とすることを特徴とするログ情報解析装置。
【請求項１２】　コンピュータの稼動状況を示すログ情報が記録されたログ情報記録部から該ログ情報が記録された時刻と所定の文字列とを抽出し、所定の時刻形式に変換された時刻と所定のフォーマットに変換された所定の文字列とを含む汎用ログを作成する汎用ログ作成ステップと、
前記汎用ログ作成ステップで作成された汎用ログを、前記所定の時刻形式に変換された時刻の順に汎用ログ記録部へ記録する汎用ログ記録ステップと、
前記汎用ログ記録部に記録された汎用ログに対応するログ情報の長さと該ログ情報中に現れる語の出現頻度とを含む特徴情報を前記ログ情報記録部に記録されたログ情報から抽出する特徴情報抽出ステップと、
前記特徴情報抽出ステップで抽出されたログ情報を図形化して前記ログ情報と共に前記所定の時刻形式に変換された時刻の順に表示する表示ステップとを備えたことを特徴とするログ情報解析方法。
【請求項１３】　請求項１２記載のログ情報解析方法において、前記特徴情報に含まれる語の出現頻度は、前記所定のフォーマットに変換された所定の文字列中に１単語が単独で出現する頻度と少なくとも２単語が連続して出現する頻度とを用いて求められることを特徴とするログ情報解析方法。
【請求項１４】　請求項１２記載のログ情報解析方法において、前記表示ステップで表示するログ情報は、前記特徴情報に含まれる語の出現頻度が所定の値を有する語または指定された所定の語を少なくとも含む文字列が識別可能に表示されることを特徴とするログ情報解析方法。
【請求項１５】　コンピュータの稼動状況を示すログ情

報を解析するログ情報解析方法を実行するコンピュータが読み出し可能なプログラムを格納した記録媒体であって、
前記ログ情報が記録されたログ情報記録部から該ログ情報が記録された時刻と所定の文字列とを抽出し、所定の時刻形式に変換された時刻と所定のフォーマットに変換された所定の文字列とを含む汎用ログを作成する汎用ログ作成ステップと、
前記汎用ログ作成ステップで作成された汎用ログを、前記所定の時刻形式に変換された時刻の順に汎用ログ記録部へ記録する汎用ログ記録ステップと、
前記汎用ログ記録部に記録された汎用ログに対応するログ情報の長さと該ログ情報中に現れる語の出現頻度とを含む特徴情報を前記ログ情報記録ステップで記録されたログ情報から抽出する特徴情報抽出ステップと、
前記特徴情報抽出ステップで抽出されたログ情報を図形化して前記ログ情報と共に前記所定の時刻形式に変換された時刻の順に表示する表示ステップとを備えたことを特徴とするログ情報解析方法を実行するコンピュータが読み出し可能なプログラムを格納した記録媒体。
【請求項１６】　請求項１５記載の記録媒体において、前記特徴情報に含まれる語の出現頻度は、前記変換された所定の文字列中に１単語が単独で出現する頻度と少なくとも２単語が連続して出現する頻度とを用いて求められることを特徴とする記録媒体。
【請求項１７】　請求項１５記載の記録媒体において、前記表示ステップで表示するログ情報は、前記特徴情報に含まれる語の出現頻度が所定の値を有する語または指定された所定の語を少なくとも含む文字列が識別可能に表示されることを特徴とする記録媒体。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、ログ情報解析装置、方法および記録媒体に関し、特にコンピュータ・システムの運用管理に用いられるログ情報解析装置、方法および記録媒体に関する。
【０００２】
【従来の技術】コンピュータ・システムの運用管理、特に障害分析または異常現象の発見等を行うためには定期的なログ情報の調査、監視または解析等の作業が重要である。特に、近年ではコンピュータ・システムへの不正侵入を検知するためにログ情報の解析作業の重要性は極めて増大している。
【０００３】しかし、従来のログ情報の解析作業は、システム管理者が既存のエディタ等を用いて膨大なログ情報を有するログファイルの閲覧を行い、独自のスクリプトを使用して特定のログを抽出していたため、極めて時間を要する作業となっていた。特に、ログ情報は種々のデータ形式を有している上にファイル・システム上に偏在しているという特徴があるため、ログ情報を解析する前の調査時における作業負担が極めて重いという問題があった。
【０００４】さらに、ログ情報は膨大な量を有しているため、システム管理者が注目するべきログ情報を抽出することは極めて困難である。既知の異常を示すログ情報の抽出は既存のキーワード検索命令を利用することにより可能ではあるが、ログ情報の中には既知の異常ではない異常を示すログ情報が含まれている場合もある。このような既知の異常ではない異常を示すログ情報を抽出することは極めて困難であるという問題があった。
【０００５】ログ情報は文字による記録であるため、システム管理者が文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することは極めて困難であるという問題があった。
【０００６】例えば、コンピュータ・システムへの不正侵入を検知する侵入検知システム（Intrusion Detection System：ＩＤＳ）では、監視する不正侵入のパターン・データをシグネチャとして作成している。新しい不正侵入の方法が発見されるたびに新しい不正侵入のパターン・データをシグネチャに登録することにより不正侵入に対応している。この結果、既知の登録されているパターン・データとは異なる不正侵入が試みられた場合には、必ずしも有効に対応することができないという問題があった。膨大なログ情報の中から警報を抽出し、さらにその警報の中から誤報を切り捨てて注目するべき警報を迅速に把握することは極めて困難であるという問題があった。
【０００７】
【発明が解決しようとする課題】上述のように、従来のログ情報の解析作業は極めて時間を要する作業であり、ログ情報の種々のデータ形式とファイル・システム上の偏在性という特徴のため作業負担が極めて重いという問題があった。さらに、膨大なログ情報の中から既知の異常ではない異常を示すログ情報を抽出することは極めて困難であるという問題があった。システム管理者が文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することは極めて困難であるという問題があった。
【０００８】そこで、本発明の目的は、上記問題を解決するためになされたものであり、ログ情報の解析作業に要する時間を減少させ、種々のデータ形式とファイル・システム上の偏在性とを有するログ情報を統合し、既知の異常ではない異常を示すログ情報を抽出することができ、システム管理者が文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することができるログ情報解析装置、方法および記録媒体を提供することにある。
【０００９】
【課題を解決するための手段】請求項１に記載された発明のログ情報解析装置は、コンピュータの稼動状況を示

すログ情報が記録されたログ情報記録部と、前記ログ情報記録部に記録されたログ情報から該ログ情報が記録された時刻と所定の文字列とを抽出し、所定の時刻形式に変換された時刻と所定のフォーマットに変換された所定の文字列とを含む汎用ログを作成する汎用ログ作成手段と、前記汎用ログ作成手段により作成された汎用ログを、前記所定の時刻形式に変換された時刻の順に記録する汎用ログ記録手段と、前記汎用ログ記録手段により記録された汎用ログを有する汎用ログ記録部と、前記汎用ログ記録部に記録された汎用ログに対応するログ情報の長さと該ログ情報中に現れる語の出現頻度とを含む特徴情報を前記ログ情報記録部から抽出する特徴情報抽出手段と、前記特徴情報抽出手段により抽出されたログ情報を図形化して前記ログ情報と共に前記所定の時刻形式に変換された時刻の順に表示する表示手段を備えたものである。

【００１０】請求項２に記載された発明のログ情報解析装置は、請求項１において、前記特徴情報に含まれる語の出現頻度は、前記所定のフォーマットに変換された所定の文字列中に１単語が単独で出現する頻度と少なくとも２単語が連続して出現する頻度とを用いて求めることができる。

【００１１】請求項３に記載された発明のログ情報解析装置は、請求項１において、前記表示手段が表示するログ情報は、前記特徴情報に含まれる語の出現頻度が所定の値を有する語または指定された所定の語を少なくとも含む文字列が識別可能に表示されることができる。

【００１２】請求項４に記載された発明のログ情報解析装置は、請求項１ないし３のいずれかにおいて、前記表示手段は、前記所定のフォーマットに変換された所定の文字列が所定の種類に該当するログ情報を選択して表示することができる。

【００１３】請求項５に記載された発明のログ情報解析装置は、請求項１ないし３のいずれかにおいて、前記表示手段は、前記特徴情報に含まれる語の出現頻度が所定の範囲内にある語を含むログ情報を選択して表示することができる。

【００１４】請求項６に記載された発明のログ情報解析装置は、請求項１ないし３のいずれかにおいて、前記表示手段は、前記ログ情報の長さが所定の範囲内にあるログ情報を選択して表示することができる。

【００１５】請求項７に記載された発明のログ情報解析装置は、請求項１ないし３のいずれかにおいて、前記表示手段は、前記所定の時刻形式に変換された時刻が所定の範囲内にあるログ情報を選択して表示することができる。

【００１６】請求項８に記載された発明のログ情報解析装置は、請求項１ないし３のいずれかにおいて、前記表示手段は、所定の文字列を有するログ情報を選択して表示することができる。

【００１７】請求項９に記載された発明のログ情報解析装置は、請求項１ないし８のいずれかにおいて、前記汎用ログ作成手段における前記所定の時刻形式は、所定の基準となる時刻から前記ログ情報が記録された時刻までの経過時間とすることができる。

【００１８】請求項１０に記載された発明のログ情報解析装置は、請求項１ないし９のいずれかにおいて、前記表示手段における前記ログ情報の図形化は、所定の時間あたりの該ログ情報の数のヒストグラム化とすることができる。

【００１９】請求項１１に記載された発明のログ情報解析装置は、請求項１ないし９のいずれかにおいて、前記表示手段における前記ログ情報の図形化は、該ログ情報に含まれる所定の文字列の長さに対応した線とすることができる。

【００２０】請求項１２に記載された発明のログ情報解析方法は、コンピュータの稼動状況を示すログ情報が記録されたログ情報記録部から該ログ情報が記録された時刻と所定の文字列とを抽出し、所定の時刻形式に変換された時刻と所定のフォーマットに変換された所定の文字列とを含む汎用ログを作成する汎用ログ作成ステップと、前記汎用ログ作成ステップで作成された汎用ログを、前記所定の時刻形式に変換された時刻の順に汎用ログ記録部へ記録する汎用ログ記録ステップと、前記汎用ログ記録部に記録された汎用ログに対応するログ情報の長さと該ログ情報中に現れる語の出現頻度とを含む特徴情報を前記ログ情報記録部に記録されたログ情報から抽出する特徴情報抽出ステップと、前記特徴情報抽出ステップで抽出されたログ情報を図形化して前記ログ情報と共に前記所定の時刻形式に変換された時刻の順に表示する表示ステップとを備えたものである。

【００２１】請求項１３に記載された発明のログ情報解析方法は、請求項１２において、前記特徴情報に含まれる語の出現頻度は、前記所定のフォーマットに変換された所定の文字列中に１単語が単独で出現する頻度と少なくとも２単語が連続して出現する頻度とを用いて求めることができる。

【００２２】請求項１４に記載された発明のログ情報解析方法は、請求項１２において、前記表示ステップで表示するログ情報は、前記特徴情報に含まれる語の出現頻度が所定の値を有する語または指定された所定の語を少なくとも含む文字列が識別可能に表示されることができる。

【００２３】請求項１５に記載された発明の記録媒体は、コンピュータの稼動状況を示すログ情報を解析するログ情報解析方法を実行するコンピュータが読み出し可能なプログラムを格納した記録媒体であって、前記ログ情報が記録されたログ情報記録部から該ログ情報が記録された時刻と所定の文字列とを抽出し、所定の時刻形式に変換された時刻と所定のフォーマットに変換された所

定の文字列とを含む汎用ログを作成する汎用ログ作成ステップと、前記汎用ログ作成ステップで作成された汎用ログを、前記所定の時刻形式に変換された時刻の順に汎用ログ記録部へ記録する汎用ログ記録ステップと、前記汎用ログ記録部に記録された汎用ログに対応するログ情報の長さと該ログ情報中に現れる語の出現頻度とを含む特徴情報を前記ログ情報記録ステップで記録されたログ情報から抽出する特徴情報抽出ステップと、前記特徴情報抽出ステップで抽出されたログ情報を図形化して前記ログ情報と共に前記所定の時刻形式に変換された時刻の順に表示する表示ステップとを備えたログ情報解析方法を実行するコンピュータが読み出し可能なプログラムを格納した記録媒体である。

【００２４】請求項１６に記載された発明の記録媒体は、請求項１５において、前記特徴情報に含まれる語の出現頻度は、前記変換された所定の文字列中に１単語が単独で出現する頻度と少なくとも２単語が連続して出現する頻度とを用いて求めることができる。

【００２５】請求項１７に記載された発明の記録媒体は、請求項１５において、前記表示ステップで表示するログ情報は、前記特徴情報に含まれる語の出現頻度が所定の値を有する語または指定された所定の語を少なくとも含む文字列が識別可能に表示されることができる。

【００２６】

【発明の実施の形態】以下、図１を用いてまず本発明の各実施の形態に共通するログ情報解析装置の概要を示し、次に図面を参照して各実施の形態を詳細に説明する。

【００２７】図１は、本発明のログ情報解析装置に用いられるコンピュータの内部回路ブロックを示す。図１において、符号１５は本発明のログ情報解析装置の内部回路ブロック、１は本発明のログ情報解析方法等を処理する処理装置ＣＰＵ（CentralProcessing Unit）、２は内部回路ブロック１５の初期化等その他の処理に必要なデータが格納された読み出し専用記憶装置ＲＯＭ（Read Only Memory）、３はＣＰＵ１が実行するコンピュータ・プログラムまたはデータが格納された読み書き可能な記憶装置ＲＡＭ（Random Access Memory）、４はＣＰＵ１に処理結果に基づいて画像を表示するディスプレイ等の画像表示部、６はログ情報解析方法等を実行するためのコンピュータ・プログラムまたはデータ等を記録したコンピュータ読み取り可能なＣＤ－ＲＯＭ（Compact Disc - Read Only memory）等の脱着可能な記録媒体をセットする記録媒体部、８はログ情報が記録されたログファイル（ログ情報記録手段）、９はログフィアル８から所望のデータを抽出して所望の変換が行われた汎用ログが記録された汎用ログファイル（汎用ログ情報記録手段）、７はログファイル８と汎用ログファイル９とが記録されたハード・ディスク等の記録装置、５は記録媒体部６または記録装置７等と接続され入出力の制御を行う入出力制御部、１１はログ情報解析装置へ入力操作を行うマウス、キーボード等の入力操作部、１０は入力操作部１１と接続され入力制御等を行う入力制御部、１２はネットワーク（不図示）を介して外部のコンピュータ等と行なう通信を制御する通信制御部、１４は上述のＣＰＵ１、ＲＯＭ２、ＲＡＭ３、画像表示部４、入出力制御部５、入力制御部１０および通信制御部１２を接続するバスである。

【００２８】本発明のログ情報解析方法等を実行するコンピュータ・プログラムおよびデータは記録媒体部６にセットされたＣＤ－ＲＯＭ等の記録媒体に記録させておくことができる。ＣＤ－ＲＯＭまたはメモリ・カード等の記録媒体に記録された上記コンピュータ・プログラムおよびデータは、入出力制御部５を介してバス１４を通りＲＡＭ３へロードされる。ＣＰＵ１はＲＡＭ３内にロードされた上記コンピュータ・プログラムを実行することにより、入力操作部１１から入力制御部１０を介して所望の入力が行なわれ、画像表示部４にログ情報解析方法を実行中の画像を表示させることができる。

【００２９】実施の形態１．図２は、本発明の実施の形態１におけるログ情報解析方法のフローチャートを示す。以下、フローチャートの各ステップを図面を参照しながら説明する。図２に示されるように、まずログファイル８からログ情報を入力する（ステップＳ１０２）。このログ情報はコンピュータの稼動状況を示す情報である。以下では説明の都合上、不正アクセスを検知するＩＤＳ等におけるログ情報を用いて説明しているが、本発明のログ情報解析方法等が対象とするログ情報は一般的なログ情報であって、ＩＤＳ等におけるログ情報に限定されるものではない。

【００３０】次に、入力したログ情報を後述の汎用ログフォーマットに変換する（ステップＳ１０４、汎用ログ作成手段）。ログ情報は、ログ情報がログファイル８に記録された時刻と、そのログ情報の内容、例えばどこから接続されたか等の内容を有するデータ（所定の文字列）とを有している。このログ情報から時刻と他の所望のデータを抽出し、汎用ログフォーマットにしたがって汎用ログを作成する。

【００３１】図３は、本発明の実施の形態１における汎用ログフォーマットを示す。図３において、符号１６は汎用ログフォーマットの全体を示し、１７は所定の時刻形式に変換された時刻情報、１８および１９は各々空白および指定の区切り文字を含まない任意の文字列であるタグ情報１およびタグ情報２、２０は指定の区切り文字を含まない任意の文字列であるメッセージ情報である。

【００３２】図４は本発明の実施の形態１におけるログ情報から汎用ログへの変換方法のフローチャートを示し、図５（Ａ）、（Ｂ）は本発明の実施の形態１におけるログ情報から汎用ログへの変換例を示す。図５で図３と同じ符号を付した個所は同じ部分を示すため説明は省

略する。

【００３３】図４に示されるように、まず時刻情報１７とする情報をログ情報から１つ決定する（ステップＳ２０２）。例えば図５（Ａ）に示されるように、時刻情報１７とする情報をログ情報ＡＣＣＴ２９Ａの時刻２３（０３：１７：１５）と決定する。

【００３４】次に、決定された時刻２３を所定の基準となる時刻、例えば１９７０年１月１日９：００からの経過秒数（所定の時刻形式）へ変換する（ステップＳ２０４）。ここで所定の基準となる時刻は任意に設定することができる。この結果、例えば図５（Ａ）に示されるように、時刻情報１７の値（変換された時刻）は、「９５２６２５８６５（秒）」と設定される。

【００３５】ログ情報から、後述のフィルタリングまたは単語の出現頻度を抽出する際の基礎情報となる情報を決定し、所定の文字列へ変換することによりタグ情報１（１８）を作成する（ステップＳ２０６）。タグ情報２（１９）も同様に作成する（ステップＳ２０８）。例えば図５（Ａ）に示されるように、基礎となる情報をログ情報ＡＣＣＴ２９Ａの文字列「ｙｐｗｈｉｃｈ」２１と決定し、この文字列２１をそのままタグ情報２（１９）として作成する。タグ情報１（１８）にはログ情報ＡＣＣＴ２８に含まれていない情報、例えばログ情報解析装置の実行環境から得られる補完的な情報（例。「ｆｏｏ．ｃｏ．ｊｐ」）を用いることもできる。

【００３６】最後に、ログ情報から解析に必要となる情報を決定し、所定の文字列へ変換することによりメッセージ情報２０を作成する（ステップＳ２１０）。例えば図５（Ａ）に示されるように、解析に必要となる情報をログ情報ＡＣＣＴ２９Ａの文字列「ｒｏｏｔ　？」２２と文字列「０．０５　０．０１　５４０．００」２４とに決定し、これらの文字列２２と２４とを組み合わせた新たな文字列「ｒｏｏｔ　？　０．０５　０．０１　５４０．００」をメッセージ情報２０として作成する。

【００３７】図５（Ｂ）は他の変換例を示す。図５（Ｂ）に示されるように、時刻情報１７とする情報をログ情報ＳＹＳＬＯＧ２９Ｂの時刻２５（Ｏｃｔ２５　１０：４１：２５）と決定する。決定された時刻２５を所定の基準となる時刻からの経過秒数へ変換し、例えば図５（Ｂ）に示されるように、時刻情報１７の値を「９３３９１２８６５（秒）」と設定する。

【００３８】タグ情報１（１８）は、図５（Ｂ）に示されるように基礎となる情報をログ情報ＳＹＳＬＯＧ２９Ｂの文字列「ｆｏｏ．ｃｏ．ｊｐ」２６と決定し、この文字列２６をそのままタグ情報１（１８）として作成する。タグ情報２（１９）は、図５（Ｂ）に示されるように基礎となる情報をログ情報ＳＹＳＬＯＧ２９Ｂの文字列「ｉｎ．ｔｅｌｎｅｔｄ」２７と決定し、この文字列２７をそのままタグ情報２（１９）として作成する。

【００３９】メッセージ情報２０は、図５（Ｂ）に示されるように、解析に必要となる情報をログ情報ＳＹＳＬＯＧ２９Ｂの文字列「ｃｏｎｎｅｃｔ　ｆｒｏｍ　ｃｒａｃｋ．ｃｏｍ」２８と決定し、この文字列２８をそのままメッセージ情報２０として作成する。

【００４０】上述のように、ログ情報を汎用ログフォーマットという中間形式に変換することにより、ファイル・システム中に偏在していたログ情報を時刻情報１７を基準として統合することができる。

【００４１】図２のフローチャートに戻り、ステップＳ１０４に続いて、特徴情報を抽出する（特徴情報抽出手段、ステップＳ１０６）。特徴情報は、汎用ログフォーマットに変換された時刻情報１７、タグ情報１（１８）、タグ情報２（１９）またはメッセージ情報（２０）等またはログ情報から抽出する情報である。後述するように、これらの特徴情報をログ情報と共に視覚化して画像表示部４に表示することにより、システム管理者は的確かつ迅速にログ情報を把握することができる。特徴情報には、単位時間当たりのログ情報の数、タグ情報１（１８）等の種類別のログ情報の数、メッセージ情報２０の文字列の長さまたはメッセージ情報２０に含まれる単語の出現頻度等がある。上述の特徴情報は例示的なものであり、これらの例に限定されるものではない。

【００４２】図６は、本発明の実施の形態１におけるメッセージ情報２０に含まれる単語の出現頻度を求める方法を説明する。図６に示されるように、ログメッセージ３０に対してテキストマイニングによる情報抽出３３を行い、１単語が単独で出現する単語別出現頻度３４と２単語が連続して出現する二単語連結句別出現頻度３７とを得ることができる。得られた単語別頻度３４と二単語連結句別出現頻度３７とは、各々図６に示されるように、出現数と単語等との一覧により示されている。

【００４３】例えば、単語別出現頻度３４に示されるように、単語「ｆｒｏｍ」はログメッセージ３０中に６回単独で出現しており、一方、単語「ｋｙｏｔｏ」３５はログメッセージ３０中に１回単独で出現し、単語「ｓｅｎｄａｉ」３６はログメッセージ３０中に１回単独で出現している。二単語連結句別出現頻度３７に示されるように、二単語「ｃｏｎｎｅｃｔ　ｆｒｏｍ」はログメッセージ３０中に４回出現しており、一方、二単語「ｆｒｏｍ　ｋｙｏｔｏ」３８はッセージ３０中に１回単独で出現し、単語「ｆｒｏｍ　ｓｅｎｄａｉ」３９はログメッセージ３０中に１回単独で出現している。この結果、単語別出現頻度３４において最も出現頻度の低い単語「ｋｙｏｔｏ」３５および単語「ｓｅｎｄａｉ」３６と、二単語連結句別出現頻度３７において最も出現頻度の低い二単語「ｃｏｎｎｅｃｔ　ｆｒｏｍ」および二単語「ｆｒｏｍ　ｋｙｏｔｏ」３８とを含むログメッセージ３１と３２とが調査を要するログメッセージであることを推測することができる。

【００４４】一般的なログメッセージは数が多く記録さ

れているため、記録されている数が少ない、すなわち出現頻度の低いログメッセージは不正侵入等を示す可能性等が高いログメッセージであると考えることができる。本発明の実施の形態１では上述のように単語別出現頻度３４と二単語連結句別出現頻度３７とを求めることにより、不正侵入等の、既知の異常とは異なる異常を示すログ情報を抽出することができ、注目するべきログ情報をシステム管理者に提示することができる。

【００４５】図２のフローチャートに戻り、ステップＳ１０６に続いて、ログ情報を視覚化して表示する（表示手段、ステップＳ１０８）。具体的には、単位時間当たりのログ情報の数をヒストグラムとして表示し、ログ情報の文字列の長さに対応した線を用いてログ情報をアウトライン的に表示し（アウトライン表示）、並列的にログ情報の文字の情報の部分を表示する。その他、タグ情報１（１８）等別のログ情報の数を表示することもできる。

【００４６】図７は、本発明の実施の形態１におけるログ情報の視覚化表示を示す。図７において、符号４０はログ情報の視覚化表示の全体を示し、ログ情報は縦軸方向の矢印（時間軸）で示されるように記録された時刻の順に上から下へ表示されている。符号４１は注目しているログのタグ情報１（１８）等の頻度を示すタグ格子領域またはタグ種別表示領域を示す。このタグ格子領域４１は図７では明示されていないが、最上部が青色で、以下黄色になり、最下部になるほど赤色で表示されている。この色分けは、後述する他の領域４２等における各種の表示において、出現頻度が多いものほど青色で表示され、出現頻度が中程度であるものは黄色で表示され、出現頻度が低いものほど注意を引くために赤色で表示することを示すための領域である。上述の色分けは例示的なものであって、出現頻度の区別ができる色分けであれば任意の色分けを用いることができる。

【００４７】図７において、符号４２はログ情報の数の時間分布を表示する時間別頻度領域または時刻情報表示領域である。時間別頻度領域４２は、時間軸方向に曜日に分かれた週格子４２ａと時間に分かれた日格子４２ｂとを有しており、横軸方向にログ情報の数が示された単位時間別のヒストグラムを表示するヒストグラム領域４２ｃを有している。ヒストグラム領域４２ｃは時間軸順に上から下へ等間隔で横線４２e、４２ｆ等が引かれている。これらの横線４２ｅと４２ｆとの間は１２時間分の時間間隔を示している。符号４５の円内に示されるヒストグラムは所定の期間にわたってほぼ一定の数のログ情報が記録されたことを示している。週格子４２ａと日格子４２ｂとを用いることにより、領域４２ｃに表示されたログ情報のヒストグラムの曜日毎、時間毎の周期的な側面の把握をさらに容易に行うことができる。符号４９の円内に示されるヒストグラムは極めて少ない数のログ情報の記録がなされている間に、符号ＬＯＧで示される極めて膨大な数のログが記録されたことを示している。このログ（ＬＯＧ）を以下では注目ログと呼ぶ。符号４２ｄは注目ログＬＯＧが記録された時刻を示している。

【００４８】図７において、符号４３はログ情報の文字列の長さに対応した線でアウトライン的にログ情報を表示するアウトライン表示領域である。アウトライン表示領域については後述する。符号４４はアウトライン表示領域４３における注目ログＬＯＧを含む注目領域４６内の実際のログ情報を拡大して表示するログメッセージ表示領域である。注目ログＬＯＧを時刻情報表示領域４２内でマウス等の入力操作部１１によりクリック等して選択すると、注目ログＬＯＧを中心として時間軸上の前後の概要をアウトライン表示領域４３に表示することができる。アウトライン表示領域４３中の注目領域４６の部分が拡大されてログメッセージ表示領域４４に対応する実際のログ情報を表示することができる。

【００４９】図８（Ａ）、（Ｂ）は本発明の実施の形態１におけるアウトライン表示領域を説明する。図８で図７と同じ符号を付した個所は同じ部分を示すため説明は省略する。図８（Ａ）において、符号５０はログメッセージ表示領域４４におけるログメッセージを一部取り出したものであり、５０ａないし５０ｄは各ログメッセージを示す。符号５１はログメッセージ５０中の各ログメッセージ５０ａ等を文字列の長さに対応した長さを有する線として表示したアウトライン表示である。アウトライン表示５１ａないし５１ｄの各線はログメッセージ５０ａないし５０ｄに対応している。このアウトライン表示５１ａ等の長さに加えて、タグ種別表示領域４１に示されるログのタグ情報１（１８）等の頻度に応じた色分けをさらに施すことができる。図８（Ａ）では色分けの代わりにハッチング等の種類で頻度を表示している。

【００５０】図８（Ｂ）において、符号５２はログ情報の視覚化表示の全体４０からアウトライン表示領域４３とログメッセージ表示領域４４とを取り出した表示を示す。アウトライン表示領域４３における注目ログＬＯＧを含む注目領域４６内の実際のログ情報が拡大されて、ログメッセージ表示領域４４内の領域５４内に表示されている。逆に、ログメッセージ表示領域４４内の領域５４内に表示されている文字情報が、注目領域４６にログ情報の長さを基に抽象化された図形となって表示されていると見ることもできる。

【００５１】図９は、本発明の実施の形態１におけるログ情報の視覚化表示の他の例を示す。図９で図７と同じ符号を付した個所は同じ部分を示すため説明は省略する。図９において、符号７１、７２および７３はメッセージ情報２０に含まれる単語の出現頻度（図６参照）に応じて色分けして表示されたログ情報である。図９では色分け表示の代わりに斜線をかぶせて示されているが、タグ種別表示領域４１に示された頻度に応じた色分けを

施して表示することができる。上述のように、ログ情報と抽出された特徴情報とを組み合わせて図形化してシステム管理者に提示することができるため、ログ情報の解析作業に要する時間を減少させることができ、システム管理者は文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することができる。

【００５２】以上より、実施の形態１によれば、種々のデータ形式により記録されているログ情報を汎用ログフォーマットという中間形式に変換することにより、ファイル・システム中に偏在していたログ情報を時刻情報１７を基準として統合することができる。単語別出現頻度３４と二単語連結句別出現頻度３７とを求めることにより、不正侵入等の、既知の異常とは異なる異常を示すログ情報を抽出することができ、注目するべきログ情報をシステム管理者に提示することができる。ログ情報と抽出された特徴情報とを組み合わせて図形化してシステム管理者に提示することができるため、ログ情報の解析作業に要する時間を減少させることができ、システム管理者は文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することができる。

【００５３】実施の形態２．図１０は、本発明の実施の形態２におけるログ情報の視覚化表示の別の例を示す。図１０で図９と同じ符号を付した個所は同じものを示すため説明は省略する。図１０において、符号７０はログメッセージ表示領域４４中のログ情報の視覚化表示、７１ないし７６はシステム管理者等から指定されたパターンマッチング用の文字列（指定された所定の語）を用いて抽出されたキーワード、７７ないし８２は上述の単語の出現頻度（図６参照）の解析により所定の範囲内の出現頻度を有するものとして指定されたキーワードである。図１０では明示されていないが、キーワード７１ないし７６は赤い背景色を有するように表示することができ、キーワード７７ないし８２は青い背景色を有するように表示することができる。図１０では色分けの変わりに、キーワード７１ないし７６は左斜線をかぶせるように表示され、キーワード７７ないしは右斜線をかぶせるように表示されている。赤い背景と青い背景とは例示的な色分けであり、他の任意の色分け、反転表示等を用いることもできる。

【００５４】以上より、実施の形態２によれば、システム管理者等から指定されたパターンマッチング用の文字列を用いて抽出されたキーワードと、上述の単語の出現頻度が所定の範囲内であるものとして指定されたキーワードとを色分けして表示することができる。このため、システム管理者は文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することができる。

【００５５】実施の形態３．図１１は、本発明の実施の形態３におけるタグ情報による情報検索機能（フィルタリング）の例を示す。図１１で図９と同じ符号を付した個所は同じものを示すため説明は省略する。図１１において、符号８５はフィルタリング処理を行う前の画面、８６は指定されたタグ情報１（１８）等の種類でフィルタリングを行った結果の画面、８７は指定されたタグ情報１（１８）の出現頻度でフィルタリングを行った結果の画面である。画面８６に示されるように、タグ情報１（１８）の種類の指定はタグ種別表示領域４１中の特定の領域８８をマウス等の入力操作部１１でクリック等して選択することにより指定することができる。図１１は、タグ情報１（１８）として「ｉｎｅｔｄ」が指定された例を示している。画面８６に示されるように、タグ情報１（１８）の出現頻度の指定はタグ種別表示領域４１中の特定の領域８９をマウス等の入力操作部１１でドラッグ等して選択することにより指定することができる。

【００５６】以上より、実施の形態３によれば、マウス等の入力操作部１１でタグ情報１（１８）等の種類またはタグ情報１（１８）の出現頻度を選択することにより、容易にログ情報のフィルタリングを行うことができる。

【００５７】実施の形態４．図１２は、本発明の実施の形態４におけるメッセージ長によるフィルタリングの例を示す。図１２で図９と同じ符号を付した個所は同じものを示すため説明は省略する。図１２において、符号９０はフィルタリング処理を行う前の画面、９２、９４は指定された閾値のメッセージ長でフィルタリングを行った結果の画面である。画面９２に示されるように、メッセージ長の閾値はアウトライン表示領域４３内でメッセージ長閾値線９１を横軸上でマウス等の入力操作部１１を利用して左右に移動させることにより指定することができる。画面９２では、注目領域９３中でメッセージ長閾値線９１より長いメッセージ長が抽出された例を示している。画面９４では、注目領域９６中でメッセージ長閾値線９５より短いメッセージ長が抽出された例を示している。

【００５８】以上より、実施の形態４によれば、アウトライン表示領域４３内でメッセージ長閾値線９１を横軸上でマウス等の入力操作部１１を利用して左右に移動させることにより、容易に所望のメッセージ長より長いかまたは短いメッセージ長のログ情報のフィルタリングを行うことができる。

【００５９】実施の形態５．図１３は、本発明の実施の形態５における時間帯によるフィルタリングの例を示す。図１３で図９と同じ符号を付した個所は同じものを示すため説明は省略する。図１３において、符号１００はフィルタリング処理を行う前の画面、１０２は指定された時間帯でフィルタリングを行った結果の画面である。画面１００は、注目ログＬＯＧを中心とした前後の時間帯をアウトライン表示領域４３内の領域１０１に示している。この状態で、例えば時刻情報表示領域４２内

の注目ログＬＯＧより時間軸上で下の位置をマウス等の入力操作部１１でクリック等することにより、画面１０２に示されるように注目ログＬＯＧより後の時間帯（アウトライン領域４３内の領域１０５）を表示することができる。時間軸上で上の位置をクリック等した場合は、注目ログＬＯＧより前の時間帯を表示することができる。

【００６０】以上より、実施の形態５によれば、時刻情報表示領域４２内の注目ログＬＯＧより時間軸上で上または下の位置をマウス等の入力操作部１１でクリック等することにより、注目ログＬＯＧより前または後の時間帯を表示することができる。

【００６１】実施の形態６．図１４は、本発明の実施の形態６における単語によるフィルタリングの例を示す。図１４で図９と同じ符号を付した個所は同じものを示すため説明は省略する。図１４において、符号１１０はフィルタリング処理を行う前の画面、１１２は指定された単語でフィルタリングを行った結果の画面である。画面１００で例えば単語「ｚｅｔａｋａ」をマウス等の入力操作部１１でドラッグ等して選択することにより、画面１１２に示されるように単語「ｚｅｔａｋａ」を含むログ情報のみを表示することができる。選択される単語は１個のみではなく複数の単語を指定することもできる。この場合、複数の単語はすべて含むか（論理積）または少なくとも１個を含むか（論理和）という指定をすることもできる。

【００６２】以上より、実施の形態６によれば、所望の単語をマウス等の入力操作部１１でドラッグ等して選択することにより、選択された単語を含むログ情報のみを表示することができる。

【００６３】実施の形態７．上述した各実施の形態の機能を実現するコンピュータ・プログラムを記録した記録媒体を本発明のログ情報解析装置に供給し、そのログ情報解析装置のＣＰＵ１が記録媒体部６等にセットされた着脱可能な記録媒体に格納されたコンピュータ・プログラムを読み取り実行することによっても、本発明の目的が達成されることは言うまでもない。この場合、上述の記録媒体から読み取られたコンピュータ・プログラム自体が本発明のログ情報解析装置の新規な機能を実現することになり、そのコンピュータ・プログラムを記録した記録媒体は本発明を構成することになる。コンピュータ・プログラムを記録した記録媒体としては、例えば、ＣＤ－ＲＯＭ、ＭＯ、メモリカード、光ディスク、フロッピー（登録商標）ディスク、ハードディスク、ＲＯＭ等を用いることができる。

【００６４】以上より、実施の形態７によれば、上述した各実施の形態の機能を実現するコンピュータ・プログラムを記録した記録媒体を本発明のログ情報解析装置に供給し、そのログ情報解析装置のＣＰＵ１が記録媒体に格納されたコンピュータ・プログラムを読み取り実行することによっても、本発明の目的を達成することができる。

【００６５】

【発明の効果】以上説明したように、本発明のログ情報解析装置、方法および記録媒体によれば、ログ情報を汎用ログフォーマットに変換し、単語別出現頻度３４と二単語連結句別出現頻度３７とを求め、ログ情報と抽出された特徴情報とを組み合わせて図形化してシステム管理者に提示することにより、ログ情報の解析作業に要する時間を減少させ、種々のデータ形式とファイル・システム上の偏在性とを有するログ情報を統合し、既知の異常ではない異常を示すログ情報を抽出することができ、システム管理者が文字による膨大な量のログ情報の中から注目するべきログ情報を迅速に把握することができるログ情報解析装置、方法および記録媒体を提供することができる。

【図面の簡単な説明】

【図１】　本発明のログ情報解析装置に用いられるコンピュータの内部回路ブロックを示す図である。

【図２】　本発明の実施の形態１におけるログ情報解析方法を示すフローチャートである。

【図３】　本発明の実施の形態１における汎用ログフォーマットを示す図である。

【図４】　本発明の実施の形態１におけるログ情報から汎用ログへの変換方法を示すフローチャートである。

【図５】　本発明の実施の形態１におけるログ情報から汎用ログへの変換例を示す図である。

【図６】　本発明の実施の形態１におけるメッセージ情報２０に含まれる単語の出現頻度を求める方法を説明する図である。

【図７】　本発明の実施の形態１におけるログ情報の視覚化表示を示す図である。

【図８】　本発明の実施の形態１におけるアウトライン表示領域を説明する図である。

【図９】　本発明の実施の形態１におけるログ情報の視覚化表示の他の例を示す図である。

【図１０】　本発明の実施の形態２におけるログ情報の視覚化表示の別の例を示す図である。

【図１１】　本発明の実施の形態３におけるタグ情報によるフィルタリングの例を示す図である。

【図１２】　本発明の実施の形態４におけるメッセージ長によるフィルタリングの例を示す図である。

【図１３】　本発明の実施の形態５における時間帯によるフィルタリングの例を示す図である。

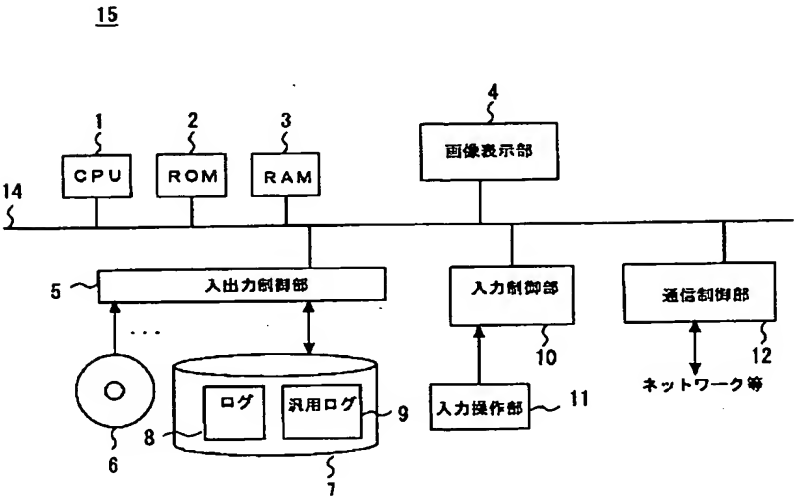
【図１４】　本発明の実施の形態６における単語によるフィルタリングの例を示す図である。

【符号の説明】

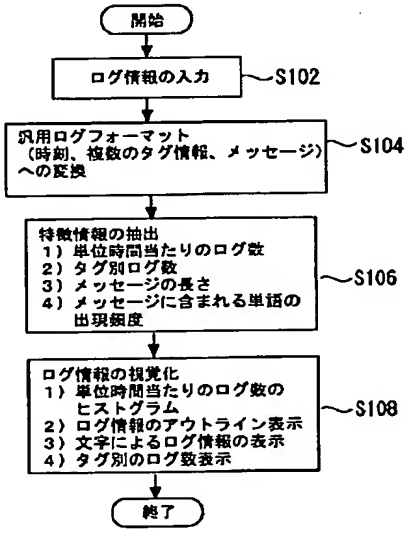
１　ＣＰＵ、　２　ＲＯＭ、　３　ＲＡＭ、　４　画像表示部、　５　入出力制御部、　６　記録媒体部、　７　記録装置、　８　ログ情報、　９　汎用ログ情報、

１０　入力制御部、　１１　入力操作部、　１２　通信制御部、　１３ネットワーク、　１４　バス、　１６　汎用ログフォーマットの全体、　１７時刻情報、　１８　タグ情報１、　１９　タグ情報２、　２０　メッセージ情報、　２１，２２，２３，２４，２５，２６，２７　文字列、　２９Ａ，２９Ｂログ情報、　３０　ログメッセージ、　３１，３２　ログメッセージ、　３３テキストマイニングによる情報抽出、　３４　単語別出現頻度、　３５，３６単語、　３７　二単語連結句別出現頻度、　３８，３９　二単語、　４１　タグ格子領域またはタグ種別表示領域、　４２　時間別頻度領域または時刻情報表示領域、　４２ａ　週格子、　４２ｂ　日格子、　４２ｃヒストグラム領域、４２ｄ　注目ログＬＯＧが記録された時刻、　４２ｅ，４２ｆ　横線、　４３　アウトライン表示領域、　４４　ログメッセージ表示領域、　４５，４９　円、４６，９３，９６　注目領域、　５０　ログメッセージの一部、　５０ａ，５０ｂ，５０ｃ，５０ｄ　ログメッセージ、　５１，５１ａ，５１ｂ　アウトライン表示、　５２　アウトライン表示領域４３とログメッセージ表示領域４４とを取り出した表示、　５４　領域、　７１，７２，７３　ログ情報、　７０　ログ情報の視覚化表示、　７１，７２，７３，７４，７５，７６　パターンマッチング用の文字列を用いて抽出されたキーワード、　７７、７８，７９，８０，８１，８２　指定されたキーワード、　８５，９０，１００，１１０　フィルタリング処理を行う前の画面、　８６　指定されたタグ情報１（１８）等の種類でフィルタリングを行った結果の画面、　８７　指定されたタグ情報１（１８）の出現頻度でフィルタリングを行った結果の画面、　９１，９５　メッセージ長閾値線、　９２，９４　指定された閾値のメッセージ長でフィルタリングを行った結果の画面、　１０２　指定された時間帯でフィルタリングを行った結果の画面、１０１，１０５　領域、　１１２　指定された単語でフィルタリングを行った結果の画面。

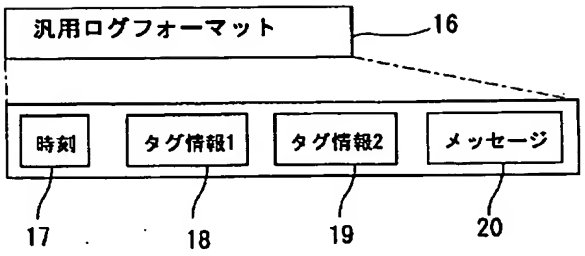
【図１】



【図２】

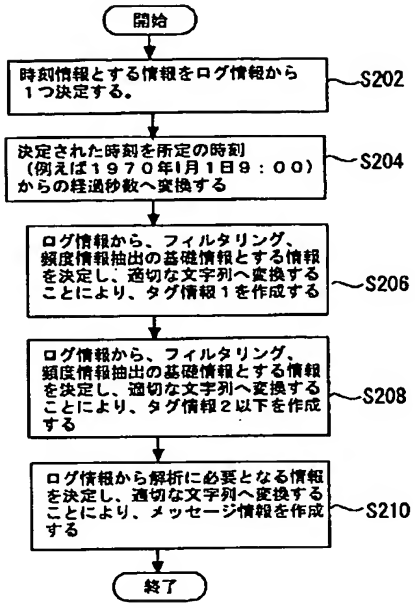


【図３】

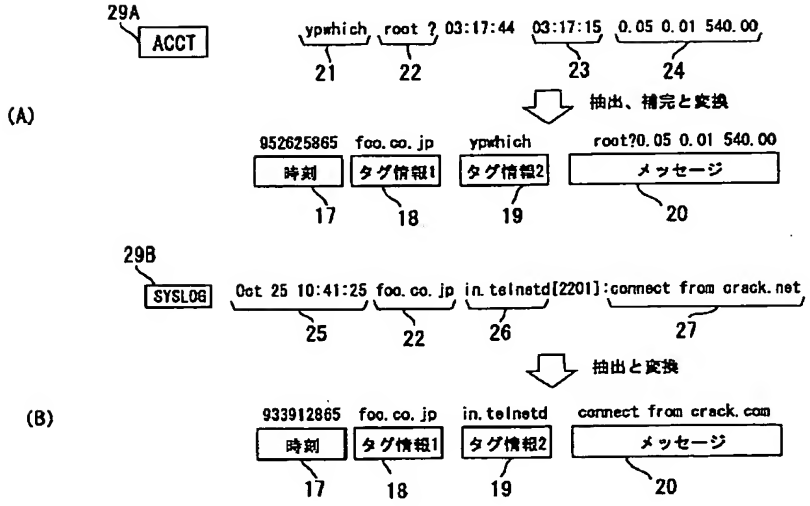


17:1970年１月１日9:00からの経過秒数
18:空白および指定の区切り文字以外の任意の文字列
19:空白および指定の区切り文字以外の任意の文字列
20:指定の区切り文字以外の任意の文字列

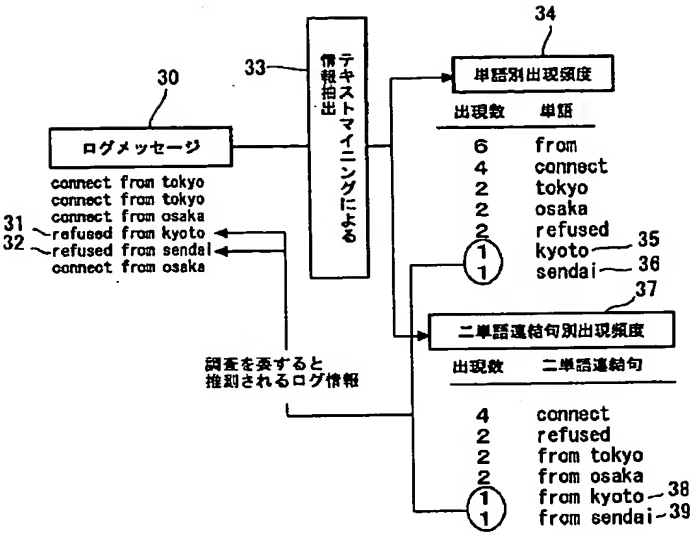
【図４】



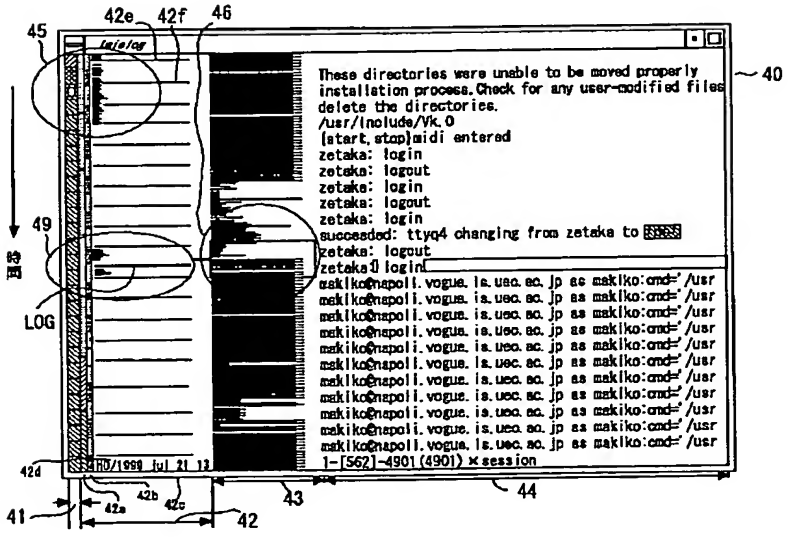
【図５】



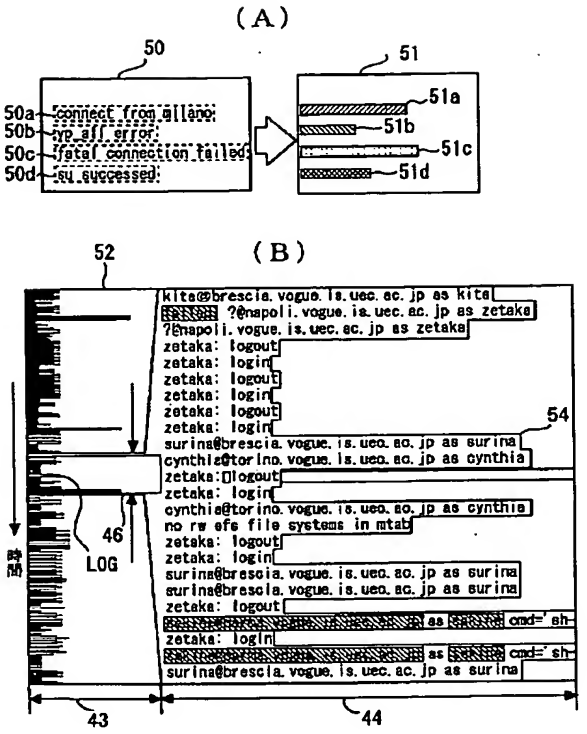
【図6】


34
33
テキストマイニングによる情報抽出
30
ログメッセージ
connect from tokyo
connect from tokyo
connect from osaka
31 refused from kyoto
32 refused from sendai
connect from osaka
単語別出現頻度
出現数 単語
6 from
4 connect
2 tokyo
2 osaka
2 refused
1 kyoto 35
1 sendai 36
37
二単語連結句別出現頻度
出現数 二単語連結句
4 connect
2 refused
2 from tokyo
2 from osaka
1 from kyoto 38
1 from sendai 39
調査を実すると
推測されるログ情報

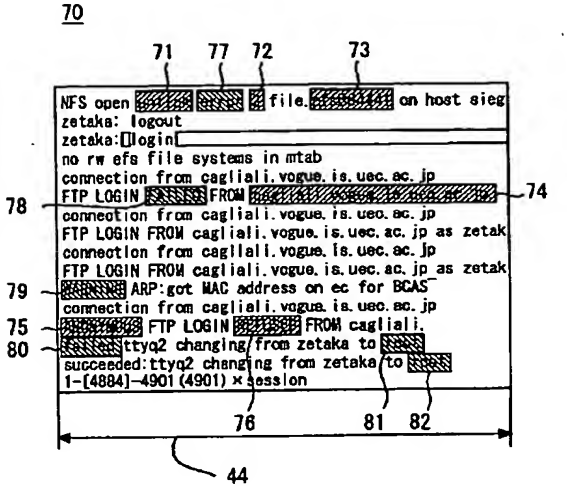
【図7】


45
42e
42f
46
40
These directories were unable to be moved properly
installation process. Check for any user-modified files
delete the directories.
/usr/include/Vk.O
(start,stop)midi entered
zetaka: login
zetaka: logout
zetaka: login
zetaka: logout
zetaka: login
succesded: ttyq4 changing from zetaka to
zetaka: logout
zetaka: login
makiko@napoli.vogue.is.uec.ac.jp as makiko:cmd='/usr
49
時間
LOG
42d
42b
42c
42a
41
42
43
44
1-[562]-4901(4901) ×session

【図8】

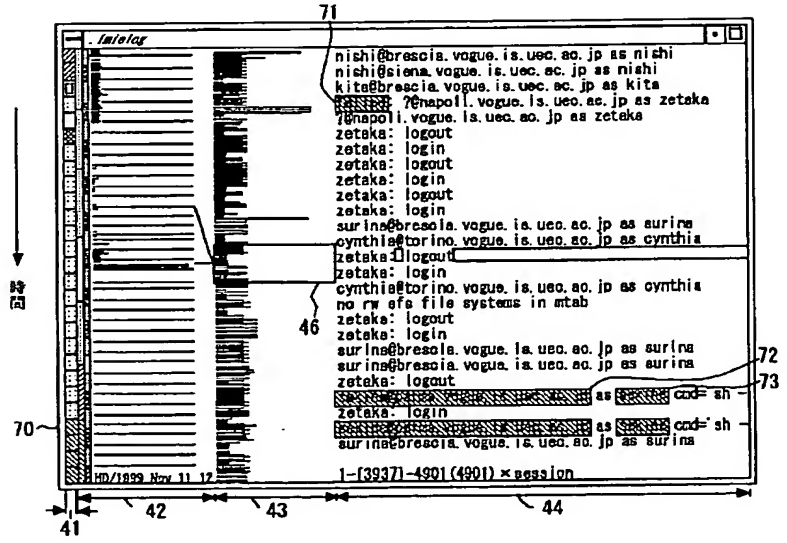


【図10】

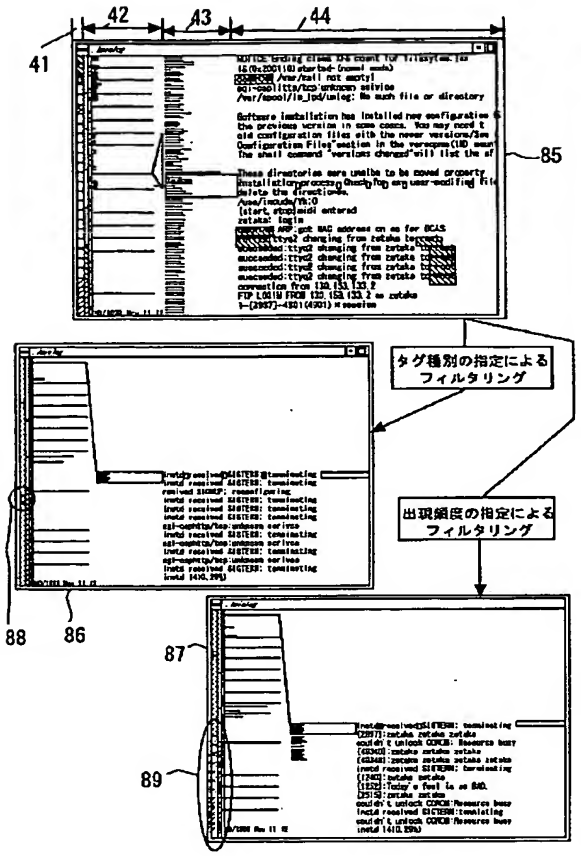


71, 72, 73, 74, 75, 76:パターンマッチングにより抽出されたキーワード
77, 78, 79, 80, 81, 82:頻度解析により抽出されたキーワード

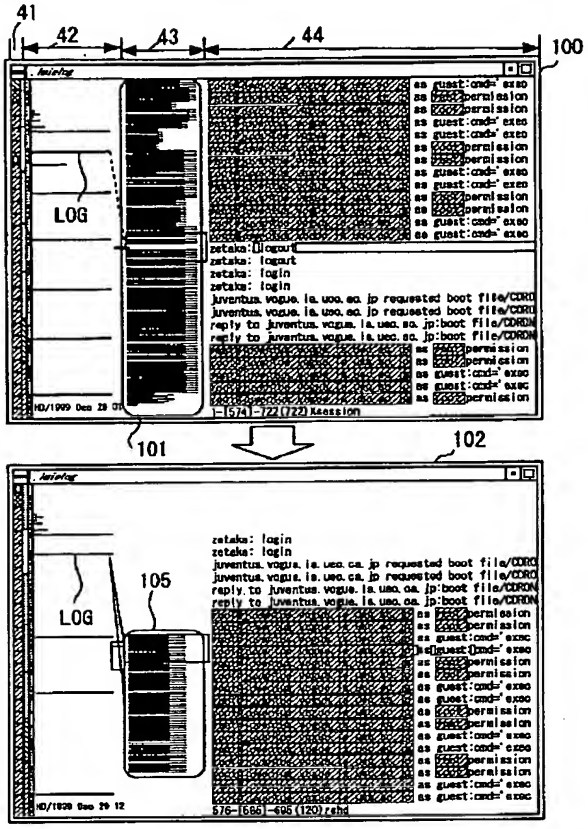
【図9】



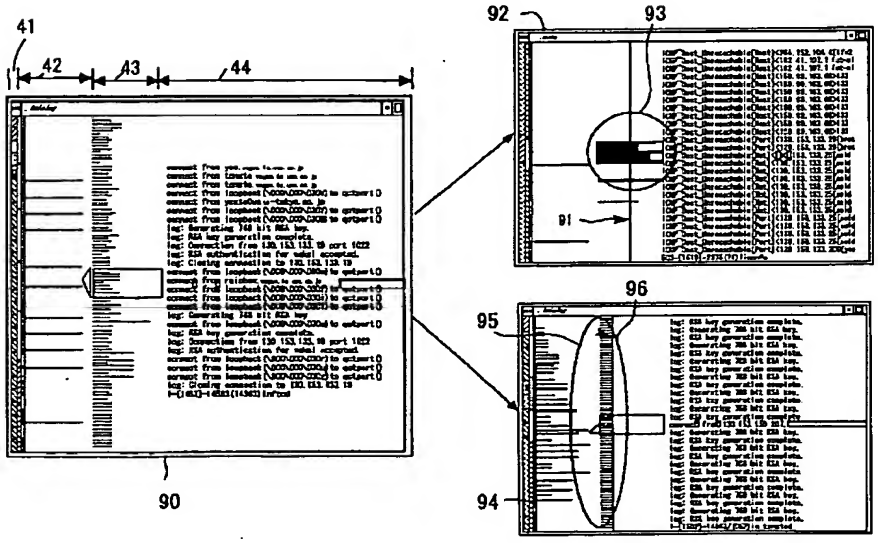
【図１１】


41
42
43
44
85
タグ種別の指定による
フィルタリング
出現頻度の指定による
フィルタリング
88
86
87
89

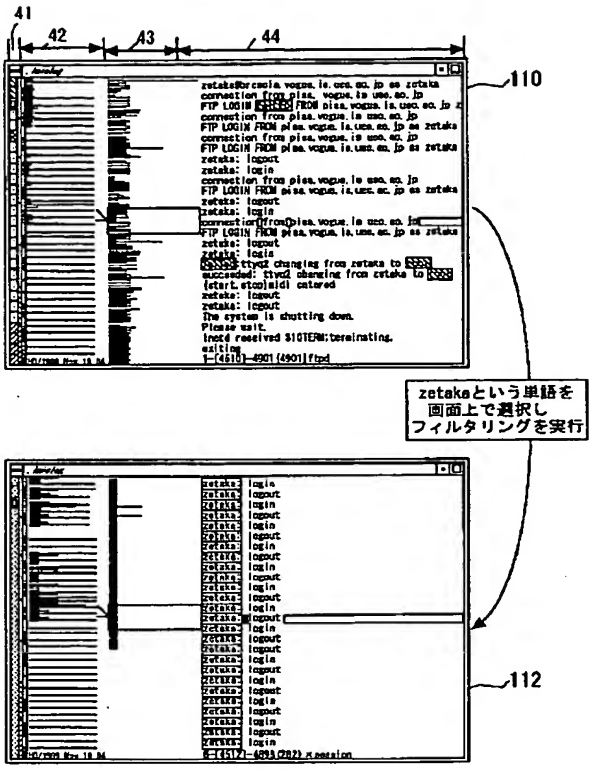
【図１３】


41
42
43
44
100
LOG
101
102
105
LOG

【図１２】


41
42
43
44
90
92
93
91
96
95
94

【図１４】



フロントページの続き

(72)発明者　岡田　幹夫
神奈川県横浜市鶴見区江ヶ崎町４番１号
東京電力株式会社システム研究所内

Ｆターム(参考)　5B042 MA08 MA11 MA14 MC15 MC23
MC28 MC35 MC40 NN04 NN09